Draft Genome Sequence of a Staphylococcus aureus Isolate Taken from the Blood of a Preterm Neonatal Blood Sepsis Patient

genomeA
Journals.ASM.org

Draft Genome Sequence of a *Staphylococcus aureus* Isolate Taken from the Blood of a Preterm Neonatal Blood Sepsis Patient

K. A. Kropp,[a] A. Lucid,[a] J. Carroll,[b] V. Belgrudov,[c] P. Walsh,[b] B. Kelly,[b] K. Templeton,[d,e] C. Smith,[d] P. Dickinson,[d] A. O'Driscoll,[c] P. Ghazal,[d] R. D. Sleator[a]

Department of Biological Sciences, Cork Institute of Technology, Bishopstown, Cork, Ireland[a]; NSilico, Cork, Ireland[b]; Department of Computing, Cork Institute of Technology, Bishopstown, Cork, Ireland[c]; Division of Pathway Medicine, University of Edinburgh, Edinburgh, United Kingdom[d]; Microbiological Diagnostic Unit, Royal Infirmary, University of Edinburgh, Edinburgh, United Kingdom[e]

Herein, we report the draft genome sequence of *Staphylococcus aureus* ED-NGS-1006, cultivated from a blood sample taken from a neonatal sepsis patient at the Royal Infirmary in Edinburgh, Scotland, United Kingdom.

Received 11 August 2014 **Accepted** 20 August 2014 **Published** 11 September 2014

Citation Kropp KA, Lucid A, Carroll J, Belgrudov V, Walsh P, Kelly B, Templeton K, Smith C, Dickinson P, O'Driscoll A, Ghazal P, Sleator RD. 2014. Draft genome sequence of a *Staphylococcus aureus* isolate taken from the blood of a preterm neonatal blood sepsis patient. Genome Announc. 2(5):e00906-14. doi:10.1128/genomeA.00906-14.

Address correspondence to: R. D. Sleator, roy.sleator@cit.ie.

Staphylococcus aureus is a Gram-positive, clinically important pathogen, and it is one of the leading causes of blood sepsis (1–3). Preterm neonates are a highly susceptible patient group for bacterial infections (3–5), and rapid detection of blood sepsis and identification of the causative agent are critical to enable proper treatment (6–8). The ClouDx-i project aims to extend current knowledge of circulating pathogenic strains linked with blood sepsis in neonates to help inform the development of new and improved molecular diagnostics. Herein, we present the draft genome of a *Staphylococcus aureus* strain, isolated from a preterm neonate at the Royal Infirmary in Edinburgh in 2013. Positivity for blood sepsis and species were confirmed by classical microbiological identification and characterization techniques.

The isolate was grown overnight at 37°C on Luria broth (LB) agar, and genomic DNA was isolated using Qiagen genomic tips (Venlo, Limburg, Netherlands). Genomic DNA was fragmented (fragments 2 to 10 kb) using sonication and a non-size selected genome library was produced, using the Nextera mate pair kit (Illumina, San Diego, CA). This library was then sequenced on an Illumina MiSeq using MiSeq Reagent kit v3. Genomic sequence assembly, analysis, and automated reporting was carried out using Simplicity (9). This approach produced 2,971,922 total reads, resulting in an average 223-fold coverage. The average G+C content was 32.93%. For sequence assembly, a *de novo* assembly pipeline based on the Spades 3.10 assembly tool was used with k-mers K21, K33, K55, K77, K99, and K127 nucleotides in length, resulting in a total of 186 contigs, of which 11 were >1,000 bp representing 97.23% of total sequence information, with the largest contig being 1,056,944 bp in size. Post assembly processing was performed by Spades and only scaffolds >1,000 bp were considered when estimating genome length as 2,814,370 bp. We annotated the contigs with Prokka (10) and used the identified 16S rRNA gene to confirm the species as *Staphylococcus aureus*. A scaffold of the genome was produced with Contiguator2 and we identified the closest related strains by BLASTing the scaffold, returning strains *S. aureus* spp. *aureus* MRSA252 and Z172 as closely related but not identical isolates, evident by many small insertions and deletions in the genomes. The genome was then screened using Glimmer3 (11) identifying 2,629 ORFs. The predicted ORFs were compared to the Uniprot-Trembl database (12) using BLASTp, mapping 1,800 ORFs to the database. To identify potential virulence factors in the genome we further compared it using BLASTp to a local database built from the VFDB (13) and Victors databases. A 75% amino-acid sequence identity cut-off was used while only considering alignments longer than 100 amino-acids, identifying 84 hits.

Samples were handled in accordance with local ethical approval by the ethics committees of the NHS Lothian SAHSC Bioresource and NHS R&D office, Project ID: 2011/R/NE/01 and the HSS BioResource RequestID: 13/ES/0126.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. JPWO00000000. The version described in this paper is version JPWO01000000.

ACKNOWLEDGMENT

This work was supported by the ClouDx-i IAPP EU FP7 project, coordinated by R. D. Sleator.

REFERENCES

1. **Foster TJ, Geoghegan JA, Ganesh VK, Höök M.** 2014. Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. Nat. Rev. Microbiol. **12**:49–62. http://dx.doi.org/10.1038/nrmicro3161.
2. **Grumann D, Nübel U, Bröker BM.** 2014. *Staphylococcus aureus* toxins—their functions and genetics. Infect. Genet. Evol. **21**:583–592. http://dx.doi.org/10.1016/j.meegid.2013.03.013.
3. **Kaufman D, Fairchild KD.** 2004. Clinical microbiology of bacterial and fungal sepsis in very-low-birth-weight infants. Clin. Microbiol. Rev. **17**:638–680. http://dx.doi.org/10.1128/CMR.17.3.638-680.2004.
4. **Shah BA, Padbury JF.** 2014. Neonatal sepsis: an old problem with new insights. Virulence **5**:170–178. http://dx.doi.org/10.4161/viru.26906.
5. **Ghazal P, Dickinson P, Smith CL.** 2013. Early life response to infection. Curr. Opin. Infect. Dis. **26**:213–218. http://dx.doi.org/10.1097/QCO.0b013e32835fb8bf.
6. **Labib AZ, Mahmoud AB, Eissa N, El Gendy FM, Soliman MA, Aly AA.**

2013. Early diagnosis of neonatal sepsis: a molecular approach and detection of diagnostic markers versus conventional blood culture. Int. J. Microbiol. Res. **4:**77–85.

7. **Mancini N, Carletti S, Ghidoli N, Cichero P, Burioni R, Clementi M.** 2010. The era of molecular and other non-culture-based methods in diagnosis of sepsis. Clin. Microbiol. Rev. **23:**235–251. http://dx.doi.org/10.1128/CMR.00043-09.
8. **Sibley CD, Peirano G, Church DL.** 2012. Molecular methods for pathogen and microbial community detection and characterization: current and potential application in diagnostic microbiology. Infect. Genet. Evol. **12:**505–521. http://dx.doi.org/10.1016/j.meegid.2012.01.011.
9. **Walsh P, Carroll J, Sleator RD.** 2013. Accelerating *in silico* research with workflows: a lesson in simplicity. Comput. Biol. Med. **43:**2028–2035. http://dx.doi.org/10.1016/j.compbiomed.2013.09.011.
10. **Seemann T.** 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30:**2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.
11. **Delcher AL, Harmon D, Kasif S, White O, Salzberg SL.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641. http://dx.doi.org/10.1093/nar/27.23.4636.
12. **Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M.** 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. **31:**365–370. http://dx.doi.org/10.1093/nar/gkg095.
13. **Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q.** 2005. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. **33:**D325–D328. http://dx.doi.org/10.1093/nar/gki008.